

Lund 1974

REPORT 7402
JANUARY 1974

Identification
of Industrial
Process Dynamics

IVAR GUSTAVSSON

LTH

Division of Automatic Control • Lund Institute of Technology

IDENTIFICATION OF INDUSTRIAL PROCESS DYNAMICS

av

Ivar Gustavsson
tekn.lic., Sm

Akademisk avhandling som för avläggande av teknisk doktorsexamen vid tekniska fakulteten vid universitetet i Lund kommer att offentligen försvaras i sal M:A, Maskinhuset, Lunds Tekniska Högskola, fredagen den 8 mars 1974 kl. 10.15.

To my parents



AV Centralen
Lund 1974

IDENTIFICATION OF INDUSTRIAL PROCESS DYNAMICS.

Most of the existing control theory requires that a mathematical model of the process and its environment is available. The lack of suitable models is one major obstacle towards a wide-spread application of modern control design techniques to industrial processes. It is thus desirable to have efficient methods to determine models for the processes to be controlled. Model building techniques are of use not only for control purposes, but also to obtain a better knowledge of a process, or to check if a certain process is performing according to the specifications. An increasing understanding of the processes involved is very often a goal in itself for systems both in industry, biology, economy etc.

Model Building Techniques.

Different model building techniques can be distinguished:

- o derivation of the model from basic physical laws and construction data,
- o estimation of models of given structure using measurements only,
- o combination of physical knowledge and estimation techniques.

Models may be derived from basic physical laws. Such models provide a good insight into the behaviour of the process and usually have a wide range of validity. Since many industrial processes are quite complicated, the models tend to be very complex. The derivation of the models is often time consuming. In many cases it may be difficult to get the required knowledge of the process. It may even be impossible, since the physics of the process are not yet fully understood. Experience shows, however, that industrial processes can often be satisfactorily controlled in steady state by simple regulators. The design of such regulators requires only approximate models.

Because of the drawbacks of modelling from physical principles, it is often necessary to determine models from measurements only. This type of modelling can be done fairly quickly. Simple models are usually obtained. A model of the disturbances can be provided. This is difficult to obtain without using process measurements. Such a model is often required for applying modern control theory. Indications of suitable approximations in models derived from physical equations can be given. Further physical knowledge of the process may be obtained. Modelling of this type sometimes provides alternative methods for measuring important physical parameters. However, in most cases it is possible to determine linear models only, which may result in a limited range of validity. Another disadvantage is, that the available a priori knowledge is seldom fully utilized.

It seems desirable to combine physical knowledge and estimation techniques in order to exploit all available knowledge and to get suitable models. In many cases it is possible to use a priori knowledge to determine the form of the relations between the process variables, i.e. the model structure, as well as to assign values to some parameters of the model. Other parameters may be unknown or very uncertain. Estimates of these parameters can then be obtained from experimental data.

Identification.

The modelling of processes from measurements is often called identification. The measured variables are usually of two kinds. Firstly the variables that can be manipulated, i.e. the inputs of the process. Secondly the process variables that are of interest, i.e. the outputs of the process. They can for example be the variables to be controlled. The concept identification is sometimes used in control literature synonymously with the statistical concept parameter estimation. However, identification is frequently used for the total modelling procedure as it will be in this thesis. It then includes:

- design of experiments,

- o determination of model structure,
- o estimation,
- o verification.

The design of experiments involves choices of all experiment conditions. For example, the variables to be measured, the instrumentation, the perturbation signal, the sampling rate, the experiment length, and so on. It must also be decided e.g. if the experiments are to be performed in open loop or in closed loop with some regulator. All these factors greatly influence the success of the identification.

By structure determination is meant the choice of a model structure, perhaps including some parameter values. The model may, for example, be static or dynamic, linear or nonlinear, deterministic or stochastic, in the time domain or in the frequency domain. It can be nonparametric, like a frequency response, or parametric, like a state space model. Important factors for the choice of model structure are the a priori information of the process, the applicability and the intended use of the model. Both physical knowledge and experimental data are used for the determination of a proper structure.

The estimation phase consists of determination of a nonparametric model or determination of the unknown parameters in a parametric model using the input output data. Also the a priori information of the process may be of value for the estimation procedure. A great number of different estimation techniques are available. The criterion for a good fit varies between the different methods, since it is not possible to give a general criterion suitable for all situations.

By verification is meant a check-up of the model and its accuracy and a test of its applicability for the intended purpose. This can preferably be done by analysing the model performance on a number of different data sets, and not only on the data used for the model fitting. If these tests indicate a poor performance, the identification must be repeated.

Development and Evaluation of Identification Methods.

Some parameter estimation techniques are very old. The well-known least squares method was developed by Gauss in the late 18th century, Gauss (1809). It is still frequently used mainly due to its simplicity. The maximum likelihood method was introduced by Fisher (1912). Different parameter estimation techniques have been studied in the statistical literature, e.g. Wold (1938), Mann and Wald (1943), Cramér (1946) and Grenander and Rosenblatt (1956). Many of the problems seem to have originated in economics. The problem of modelling economic systems from empirical data has in fact developed into a special branch, econometrics. Parameter fitting has also been frequently used in biology. In control applications non-parametric methods like frequency analysis received much attention for a long time. The models obtained by these methods are well suited for control design, using classical methods like Bode-Nichols synthesis. During the last ten years parameter estimation techniques have received more attention by control engineers due to the development of modern control theory. It has been possible to use more sophisticated control procedures based upon mathematical models. Another reason is the advent of digital computers making calculations feasible, that were earlier impossible to carry out in a reasonable time and to a reasonable cost. Many efficient identification algorithms have also been developed.

Papers about identification techniques have abounded the control literature over the last decade. Many identification problems that have only recently been recognized by control engineers have, however, actually been posed and in many cases also solved fairly generally in related topics like econometrics, statistics, and information theory. An overview of identification and parameter estimation methods, particularly used in control contexts, can be found in some of the many good survey papers that have been published, e.g. Åström and Eykhoff (1971). More references are given in Part II of this thesis. Only a few textbooks are yet available. Graupe (1972) and Eykhoff (1973) can be mentioned. Until recently

the number of applications of parametric methods and comparisons between different methods were few. For the time being, however, the number of applications is significantly increasing. A number of comparisons has also been published lately.

In this thesis several different aspects of identification are discussed. This thesis consists of four parts:

- I. Comparison of Different Methods for Identification of Industrial Processes. *Automatica*, Vol. 8, No. 2, pp 127-142, 1972. An earlier version was presented at the 2nd IFAC Symposium on Identification and Process Parameter Estimation, Prague, Gustavsson (1970).
- II. Survey of Applications of Identification in Chemical and Physical Processes. *Identification and System Parameter Estimation, Proceedings of the 3rd IFAC Symposium, The Hague/Delft, 1973* (Ed. P. Eykhoff). Invited survey paper. Also submitted to *Automatica* for possible publication.
- III. Identification of Drum Boiler Dynamics. *Identification and System Parameter Estimation, Proceedings of the 3rd IFAC Symposium, The Hague/Delft, 1973* (Ed. P. Eykhoff). Invited case study. Coauthor K. Eklund.
- IV. Identification of Linear, Multivariable Process Dynamics using Closed Loop Experiments. Report 7401, Division of Automatic Control, Lund Institute of Technology, Lund, Sweden (1974). Coauthors L. Ljung and T. Söderström.

The thesis covers many of the different phases of identification. The emphasis of the thesis is on the practical application of identification rather than on the theoretical problems. Part I is a comparison of some identification methods applied to data from different industrial processes. Part II is a survey of applications in a particular field of industrial processes. It also includes a discussion of the practical problems involved in identification. Part III is a study of identification of a drum boiler. Both linear

and nonlinear models are developed. In Part IV the identifiability of the open loop characteristics from closed loop experiments is treated. In the following the contents of each part are reviewed.

Comparison of Four Different Identification Methods.

Many identification methods have been proposed. In order to facilitate a meaningful choice between them in a practical situation it is very valuable to have comparisons between them available. One of the first extensive comparisons, using data from industrial processes, is presented in Part I. Four different methods are compared using data from a number of real processes:

- a nuclear reactor,
- a distillation column,
- a superheater,
- a paper machine.

The methods used were

- correlation and spectral analysis,
- the least squares method,
- the tally principle method,
- the maximum likelihood method.

It is difficult to compare models obtained by different identification methods and to choose model order for parametric methods in a relevant way. If the ultimate goal of the identification is the design of a control strategy, the criterion for the judgement of the models should be the closed loop performance. In practice it is seldom possible to carry out tests of many control strategies. Other ways of comparing models must therefore be used. In Part I comparisons are made between the model parameters and between transient and frequency responses of the different models. Even if the closed loop performance has not been checked, the comparisons give some in-

sight into the behaviour of different identification techniques. Different criteria for the choice of model order for parametric models are also discussed. It is demonstrated that the choice of model order must be based on a series of tests. Nevertheless, the final decision must often be based on subjective grounds. In order to compare the performance of the identification methods also on data from known processes, simulations are used.

The comparison shows that there are small differences in performance of different identification methods for data with small disturbances. Criteria like computing time or core memory requirements may determine which method is most suitable in a particular situation. On the other hand, for data from processes with a high noise to signal ratio it is necessary to choose an effective method. Good low order models for the studied processes have been obtained by the maximum likelihood method, developed by Åström and Bohlin (1966).

Survey of Applications and Practical Identification Problems.

For a control engineer going to perform model building by identification it is valuable to get references to other applications in the same field. It is also very important to recognize the practical problems that may arise. This can be done easier if a catalogue of the problems and their practical solutions is available.

Part II of this thesis consists of two sections. The first one reviews recent applications of identification to

- o chemical processes,
- o metallurgical processes,
- o paper and pulp processes,
- o glass and cement processes.

It is obvious from the review that identification has been a useful tool in this field to determine process dynamics.

This is particularly true for processes which are difficult to model from physical considerations only, like blast furnaces, paper machines and cement kilns. Valuable knowledge was often provided, not only for control purposes, but also for process design.

In the second section different practical identification problems are discussed, viz.

- o design of experiments,
- o determination of model structure,
- o identification/estimation techniques,
- o model verification,
- o data handling.

An attempt is made to summarize the experience from several applications made and to give some practical advice, which can be of value in order to get a successful modelling by identification. Only single-input single-output systems are considered. A major conclusion is that interactive program packages are very efficient tools for modelling. Problems, which have taken weeks to solve on a big computer because of the long turn around time for programs, can now be solved in hours on a medium-sized computer. The reason for the efficiency is that identification is always an iterative procedure. It is necessary to check the results and to decide what to do next. A complete automation seems impossible. Such an interactive program package has been developed, Gustavsson, Selanders and Wieslander (1973). It has been used extensively for a great number of applications. Another conclusion of the study is that it seems possible to design suitable experiments from quite simple rules of thumb as soon as some basic knowledge of the process is available. It is often not worthwhile to try to obtain optimal input sequences, sampling rates etc. One reason is that the optimal conditions require knowledge of the process model.

A Case Study - Identification of a Drum Boiler.

In Part III the results of identification of a 160 MW drum boiler - turbine unit are presented. The intention has been to highlight different aspects on all phases of identification of a full scale plant.

Three types of models are considered. Single-input single-output linear difference equation models are developed relating fuel flow and control valve position to the outputs drum pressure and active power. A multivariable, linear state space model of 5th order was derived from physical laws and construction data, Eklund (1971). Some of the parameters in it were then estimated from the measurements. The unknown parameters of a multivariable, nonlinear model of low order have also been identified. The structure was mainly obtained from physical considerations. In all cases additional insight into the physical properties of the process was gained. Though the models have not been used for control, the achieved accuracy indicates that a control system design can be successfully based on the obtained models.

The results concerning the linear models are a summary of results obtained by Eklund (1971). The new contribution is a feasibility study of nonlinear identification. It seems that the proposed method offers possibilities for nonlinear identification. The computing time required is, however, a severe limitation for the identification of multivariable and nonlinear processes. Computing times of 30 minutes on a UNIVAC 1108 to obtain a model are not rare. This can be compared with the time required to obtain a linear single-input single-output difference equation model which is less than 30 seconds. The numerical aspects of the different algorithms are briefly discussed.

It is recognized that imperfections of the data caused by non-simultaneous sampling or imperfect input signal can give rise to difficulties. This is particularly the case if the ultimate aim is to use the data for test of a model derived from physi-

cal a priori knowledge. It is then important that the measured variables are exactly the physical variables used in the model and that the instruments are properly calibrated. If the goal is just to derive a control strategy, an important factor is to measure the input and output variables in the same way as assumed by the controller. It is also recognized that design of experiments for multivariable and nonlinear systems may cause extra difficulties. The ideal experiment will depend on the structure of the particular process. A method is proposed for the exploration of several experiments simultaneously for parameter estimation. It provides the possibilities to use experiments performed at different operating conditions and can often reduce the need of multivariable experiments. The method is practically feasible as long as the model does not have too many state variables.

Identifiability of Systems Operating in Closed Loop.

There are still unsolved theoretical problems in all phases of identification. One problem, that has often been discussed, is whether it is possible to estimate the open loop characteristics of a process from measurements obtained under closed loop operation. For control system design purposes the open loop characteristics are often desired. Risks of personal injuries, process damages, loss of production or reduced efficiency may prevent open loop experiments. Other types of processes, e.g. many economical and biological processes, are inherently in closed loop.

The identifiability of the open loop characteristics of linear, multivariable systems, described by vector difference equations, and with arbitrary feedback is discussed in Part IV. The asymptotic case, i.e. when the number of samples tends to infinity, is considered. The first section gives a survey of earlier contributions treating identification of systems, operating in closed loop.

The identifiability of a system concerns the possibility to

obtain, from input output data, a model, that in some sense is a correct description of the system. Identifiability depends on the structure of the system, the set of models considered, the identification method chosen and the experiment conditions like input signal properties and feedback structure. In particular the influence of the feedback structure is studied in Part IV. It is assumed that identification methods minimizing some criterion of the prediction errors are used.

For a systematic treatment of the problem some identifiability concepts are introduced. If the identification method used gives a model that has the same input output relationship and the same noise characteristics as the system, the system is said to be System Identifiable for the chosen set of models and the experiment conditions used. If the system is system identifiable for all model structures, that include the true system transfer function and noise characteristics, the system is said to be Strongly System Identifiable for the experiment conditions used. It is sufficient for control purposes to know the transfer function. In many cases, however, it is desirable to estimate certain parameters in a given model structure. If the parameters can be estimated consistently, the system is said to be Parameter Identifiable for the chosen set of models and the experiment conditions used. Given a system identifiable process, parameter identifiability is mainly a question of suitable parametrization of the model.

At least two approaches to the identification of a system operating in closed loop are possible. One way is simply to forget the feedback and proceed as in the open loop case. Another way is to use the following indirect method. The closed loop transfer function is firstly estimated. If the regulator is exactly known, it may then be possible to determine the open loop characteristics from the closed loop transfer function.

There are significant advantages by using the first simple approach. It can be applied more generally, e.g. in the case

of manual control or when the feedback is time varying. The regulator does not have to be known, and no extra algorithms besides those used for open loop identification are necessary. It is shown in Part IV that the straightforward method has the same identifiability properties as the indirect method.

It is also shown in Part IV that if the regulator is nonlinear or time-varying or an extra perturbation is injected, then the identifiability is most often secured. This means that the presence of feedback introduces no extra difficulties concerning choice of model structure in these situations. It also means that the identifiability does not depend on the detailed structure of the regulator. If, however, the regulator is noise free, linear and time invariant and no extra perturbation is injected the system may or may not be identifiable. The identifiability depends in this case on the structures of the system, the model and the regulator. It is not possible to check the identifiability without having a detailed knowledge of these structures. Necessary and sufficient conditions for identifiability are given for single-input single-output systems with such a regulator. One possible way to secure identifiability is to inject an extra perturbation signal. In case this is not possible, it is proposed to use experiments with a time-varying regulator. This can be obtained simply by shifting between a number of linear, time invariant regulators. The number of regulators required is shown to depend only on the number of inputs and outputs.

Acknowledgements.

I wish to express my sincere gratitude to my supervisor Karl Johan Åström for his encouragement, help and guidance throughout this work. I would also like to thank all my colleagues at the Division of Automatic Control for many stimulating discussions and the very valuable team work. In particular I wish to thank my coauthors Karl Eklund, Lennart Ljung and Torsten Söderström. Thanks also to Gudrun Christensen who has typed the manuscript. This work has been supported by the Swedish Board for Technical Development under contract 733546.

References.

- Åström, K.J., and Bohlin, T. (1966)
Numerical Identification of Linear Dynamic Systems from Normal Operating Records. In Theory of Self-Adaptive Control Systems, Proceedings of the 2nd IFAC Symposium on the Theory of Self-Adaptive Control Systems, September 1965, Teddington, England (Ed. P.H. Hammond), Plenum Press, New York.
- Åström, K.J., and Eykhoff, P. (1971)
System Identification - A Survey. Automatica, Vol. 7, No. 2, pp 123-162.
- Cramér, H. (1946)
Mathematical Methods of Statistics. Princeton.
- Eklund, K. (1971)
Linear Drum Boiler - Turbine Models. Report 7117, Division of Automatic Control, Lund Institute of Technology, Lund, Sweden.
- Eykhoff, P. (1973)
System Parameter and State Estimation. (In manuscript)
- Fisher, R.A. (1912)
On an Absolute Criterion for Fitting Frequency Curves. Mess. of Math., Vol. 41, p. 155.

Gauss, K.F. (1809)

Theoria Motus Corporum Coelestium. A translation to English: Theory of the Motion of the Heavenly Bodies. Dover, New York (1963).

Graupe, D. (1972)

Identification of Systems. Von Nostrand Reinhold Company, New York.

Grenander, U., and Rosenblatt, M. (1956)

Statistical Analysis of Stationary Time Series. Almqvist and Wiksell, Uppsala.

Gustavsson, I. (1970).

Comparison of Different Methods for Identification of Linear Models for Industrial Processes. Preprints of the 2nd IFAC Symposium on Identification and Process Parameter Estimation, Prague.

Gustavsson, I., Selander, S., and Wieslander, J. (1973)

IDPAC - User's Guide. Report 7331. Division of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Mann, H.B., and Wald, A. (1943)

On the Statistical Treatment of Linear Stochastic Difference Equations. Econometrica, Vol. 11, pp 173-220.

Wold, H. (1938)

A Study in the Analysis of Stationary Time Series. Almqvist and Wiksell, Uppsala.



LT H